

UNIVERSITY OF WAIKATO

**Hamilton
New Zealand**

**Agglomeration Externalities and 1981-2006
Regional Growth in Brazil**

Valente J. Matlaba, Mark Holmes, Philip McCann and Jacques Poot

Department of Economics

Working Paper in Economics 07/12

June 2012

Corresponding Author

Valente Matlaba

Economics Department

University of Waikato

Private Bag 3105

Hamilton, New Zealand

Email: valente.matlaba@yahoo.com.br

Mark Holmes

Economics Department, University of Waikato

Email: holmesmj@waikato.ac.nz

Philip McCann

Department of Economic Geography, University of Groningen

Email: p.mccann@rug.nl

Jacques Poot

National Institute of Demographic and Economic Analysis

University of Waikato

Email: jpoot@waikato.ac.nz

Abstract

This paper focuses on manufacturing employment growth across the 26 states of Brazil. We employ the Glaeser et al. (1992) approach to identify the role played by knowledge externalities in growth and convergence. To assess robustness of the results, we compare cross-section models, dynamic panel models and pooled-periods fixed-effect models. We find that cross-section models confirm the positive impact of Porter's and Jacobs' competition externalities on growth, whereas dynamic panel models and pooled-periods fixed-effect models are consistent with the predictions of Marshall-Arrow-Romer and Porter regarding the role of specialisation in manufacturing vis-à-vis other employment. The results provide new insights into the rapid growth since 1981 in particularly the North and Centre West of Brazil.

JEL Classification

C52; O18; O31; N66; R11

Keywords

convergence
economic growth
knowledge externalities
Brazil

Acknowledgements

Earlier versions of the paper were presented at the 6th Australasian Development Economics Workshop, 10-11 June 2010, Parramatta, Australia and at the 57th North American meetings of the Regional Science Association International, 10-13 November 2010, Denver, USA. We thank the participants and an anonymous referee for helpful comments and suggestions. The first author also acknowledges the hospitality of the Faculty of Spatial Sciences at the University of Groningen during the revision of this paper.

1. Introduction

Regional income convergence is a key area of concern among economists (Barro & Sala-i-Martin, 2004). If regions with a lower real income per capita grow faster than those with a higher real income per capita, economic growth can reduce interregional income inequality within a country. A standard model that explains the role of knowledge externalities in such convergence is the Glaeser *et al.* (1992) model (see also Henderson *et al.*, 1995; Combes, 2000). Using a simple Cobb-Douglas production function, this model shows that the growth rate of employment is positively related to growth of technological knowhow and negatively related to the growth rate of wages. Assuming that wages are spatially equalised, employment growth differentials are then explained by the region-specific impact of knowledge externalities due to regional specialisation, competition and diversity, based on the theories of Marshall-Arrow-Romer (MAR), Porter and Jacobs. The Glaeser *et al.* (1992) model aims to quantify each of these effects separately.

The theories advocated by Romer (1986) and Porter (1990) argue that the concentration of specialised industries in an area tends to benefit growth in that area because knowledge externalities enhance the productivity of clustered firms.¹ In contrast, Jacobs (1969) argues that a cluster of specialised industries in a specific area can actually reduce growth because specialisation tends to inhibit competition among firms, thereby limiting regional growth potential. When competition among firms is limited, there is no place for diversity and the effects of knowledge will be impeded insofar as not reaching out to broader sectors. Thus, the expected relationship between manufacturing industry employment growth and dynamic knowledge externalities and competition is ambiguous because it depends on whether the adopted perspective is that of MAR, Porter or Jacobs (De Groot *et al.*, 2009, p. 264).

Despite the international acceptance and utilisation of the Glaeser *et al.* (1992) empirical model of the impact of these knowledge externalities, the model has not yet been applied to Brazil. The present paper therefore estimates the model for Brazil. There has been significant income convergence among Brazilian states in recent decades. This paper focuses on the role played by knowledge externalities on growth in manufacturing industry employment and, by implication, real incomes over the period 1981 to 2006. While there are many studies of economic growth, convergence and regional income inequality in Brazil, none to date has adopted the Glaeser *et al.* (1992) model. To assess robustness of the results, we employ a number of common empirical estimation approaches, based on cross-section data, panel data and pooled-periods cross-section data.

The paper is structured as follows. The following section discusses the background literature on regional income convergence in Brazil. Section 3 describes the data used and their sources. Section 4 provides a descriptive analysis of regional growth in Brazil. Section 5 reports and discusses the regression results. Lastly, section 6 provides concluding comments.

¹ These theories are summarised in Glaeser *et al.* (1992), Henderson *et al.* (1995) and De Groot *et al.* (2009).

2. Literature

Regional income convergence in Brazil has been widely studied. The literature has taken four approaches. The first approach estimates regressions models using cross-section data. These regressions utilise the growth of per capita income over a given period in a region as the dependent variable and assume that this is negatively related to initial per capita income. This negative relationship is referred to as beta convergence (Azzoni, 1999; Azzoni, 2001, 2003; Silveira-Neto & Azzoni, 2006). Other predictors of growth are investment, education and the size of labour force (Ferreira, 2000). Other regressions explain per capita income of Minimum Comparable Areas (MCAs) in terms of education, demographic indicators and the ratio of public expenditure over revenue of these areas (Rangel *et al.*, 2008). These studies find evidence in favour of absolute income convergence across Brazil's states. Additionally, Coelho & Figueiredo (2007) focus on municipalities' per capita income growth between 1970 and 2000 and explain growth in terms of a range of demographic and socioeconomic indicators (essentially education, per capita capital, and the crime rate). Brauch & Monasterio (2007) focus on income convergence by means of Exploratory Spatial Data Analysis (ESDA) of the income shares of MCAs. In an application of Quah's (1997) methodology to examine of distribution of income across Brazilian regions, Gondim *et al.* (2007) find convergence only within macro regions. Indeed, Magalhães *et al.* (2005, p. 17) point out that:

(...) it is possible to infer from the results in hand that, although some convergence among states is taking place, it seems to be more of a regional phenomenon or perhaps some type of club convergence than a global convergence process. States like São Paulo would be a dominant force in one club while the Northeast states would form a second group or club.

In other words, this study and several others (e.g. Resende, 2011, and references therein) found clubs of convergence – a situation in which rich states and poor states converge within their macro regions, but in which interregional income disparity actually increases.

Silva & Silveira Neto (2007) apply Hanson's (1998) model and evaluate the role of knowledge externalities through the estimation of manufacturing industry employment growth across Brazilian states from 1994 to 2002. Manufacturing industry employment growth is regressed against the following variables: average wage per worker, relative size of establishments, forward/backward linkages, agglomerations, manufacturing diversity, and distance. Market linkages and manufacturing diversity are both found to be positively associated with growth of manufacturing industry employment, thus confirming Jacobs' and Porter's theories, while rejecting MAR theory.

The second approach to investigate regional income convergence is to use panel data analysis. To our knowledge, the only study that uses this approach is Azzoni *et al.* (2000) who explain per capita income by geographic variables, labour force and human capital variables. They find conditional income convergence across Brazilian states.

The third approach applies time-series data analysis. For example, Barossi-Filho & Azzoni (2003) use a sample of 20 Brazilian states to study the convergence of state per capita GDP in terms of the lag of income, the national business cycle and structural breaks. After performing unit root tests they find state-level convergence within Brazilian macro-regions.

The second and third approaches described above are rather rare in the literature regarding regional income convergence in Brazil. A fourth approach is simply descriptive and measures the inequality pattern in terms of income inequality indicators or national income shares for selected time periods, rather than by means of estimating structural models. This approach finds either states' income convergence at a national level or states' income convergence only within macro-regions (Rolim, 2008; Andrade & Serra, 2001; Azzoni, 1997; Mossi *et al.*, 2003; and Gomez, 2002).

This paper builds on this literature by estimating Glaeser *et al.*'s (1992) model. The original Glaeser *et al.* (1992) model can be represented by the following equation:

$$\alpha \log \left(\frac{l_{t+1}}{l_t} \right) = -\log \left(\frac{w_{t+1}}{w_t} \right) + \log \left(\frac{A_{national,t+1}}{A_{national,t}} \right) + g(\text{specialization, competition, diversity, initial conditions}) + e_{t+1} \quad (1)$$

which can be estimated by data on city-specific employment growth by industry. In the present application we consider, due to data limitations, manufacturing employment vis-à-vis employment in other sectors. Hence, on the left side of the equation (1), l is manufacturing industry employment. Taking two points of time, and assuming $\alpha=1$ without loss of generality, the dependent variable is manufacturing industry employment growth between period t and $t+1$. This is explained in terms of national wage (w) growth (negatively), the national growth of technology (A), and by the g function which captures specialisation, competition, and diversity externalities; as well as initial conditions (e.g. human capital or other variables considered relevant by the researcher); and lastly, e_{t+1} is a residual.

Glaeser *et al.*'s model was first estimated by means of industry-by-city manufacturing employment data in the US (Glaeser *et al.* 1992; Henderson *et al.* 1995) and by manufacturing and services employment data in France (Combes, 2000) before it became more widespread. De Groot, Poot and Smit (2009) summarise the international literature covering the period 1997-2006. Based on meta-analysis – which consists of combining all empirical evidence in which at least one of the first three elements of the g function of equation (1) is tested for statistical significance – they found 322 articles that analyse dynamic externalities and that cited either Glaeser *et al.* (1992) or both Porter (1990) and Jacobs (1969).

Of these 322 articles, only 31 had a quantitative approach that was sufficiently similar to Glaeser *et al.* (1992) to permit a meta-analytic comparison and, together, these 31 articles yielded a total of 393 estimates. A summary of the significance tests conducted in this international literature is reproduced in Table 1. This table shows counts and percentages for

each result in the following classification: “negative significant”, “negative insignificant”, “positive insignificant” and “positive significant” under the following three elements of agglomeration externalities: specialisation, competition and diversity.

Table 1: Counts of Statistical Significance of Knowledge Externalities in Empirical Estimates of the Glaeser et al. (1992)-Type Models

	Specialisation		Competition		Diversity	
	Count	%	Count	%	Count	%
Negative significant	60	37	16	20	17	11
Negative insignificant	33	20	13	16	40	26
Positive insignificant	16	10	19	24	37	24
Positive significant	53	33	31	39	58	38
Total	162	100	79	100	152	100

Source: De Groot, Poot and Smit (2009, p. 269)

On balance, Table 1 indicates that there is slightly more evidence to suggest the impact of specialisation is more negative than positive. In contrast, the evidence supports on average a positive effect for competition and diversity externalities. Nonetheless, there is considerable uncertainty in the empirical evidence given the percentage of statistically insignificant results. This varies between 32 percent in the case of competition and 77 percent in the case of diversity.

3. Data

This paper makes use of two official data sources in Brazil. The first is IPEA - Instituto de Pesquisa Econômica Aplicada (Institute of Applied Economic Research).² This data source provides a variety of socio-economic data collected from public and private institutions. Regional data are available for municipalities, states and macro-regions.

We employ three types of data. First, we use labour market data, in which we extracted information on the number of employed people in each state from 1981 to 2006. There are also state-level data on the *percentage* of employed people across nine sectors of economic activity. These are: (1) agriculture & fishing; (2) commerce; (3) construction; (4) electricity, water & gas; (5) finance; (6) manufacturing; (7) mining; (8) services and (9) transportation & communications. We calculated for each state annual employment by sector by multiplying the percentage of employed people in each sector by the total number employed people of that state. Aggregated across all states, calculated employment represents between 92.22 percent and 99.21 percent of reported country-level employment between 1981 and 2006 (after imputing some values in years with missing data). Secondly, we considered the number of manufacturing sector establishments. IPEA data provide one observation on manufacturing sector establishments for the following years: 1970, 1975, 1980, 1985 and 1995. We used 1985 data to calculate the competition variable that we included in the cross-section models of manufacturing industry employment growth from 1985 to 1995 (see Table 2).

² The IPEA data can be found on the website www.ipea.gov.br.

Table 2 (at end of paper)

Thirdly, we used education data. The IPEA data present a variety of education indicators. However, there are only five indicators with annual data from 1981 to 2007. These are: the percentage of children who are behind in their schooling, the percentage of school drop-outs, the percentage of children's school attendance, the average years of schooling (for people aged 25 years and older), and the literacy rate (100 minus the percentage of illiterate people aged 15 years and older). We selected the literacy rate as the appropriate education indicator as it is a standard measure of the level of development.

However, there are two data caveats. First, there are missing data for 1991, 1994 and 2000. For these years, we interpolated the literacy rate by calculating the average between the neighbouring years for each of the years. With respect to employment, we applied the following two-step procedure: first we took the average of the distribution of the employed population across sectors between the neighbouring years for the states with missing sectoral employment data. Next, we applied those percentages to total employment in the state in each of the years listed above. This yielded for each state and each year manufacturing industry employment and employment in each of the other eight sectors. Ultimately, we created a complete time series of employment by sector and state for a period of 26 years.

The second data caveat is the level of data aggregation. The level of disaggregation of the data by region and by sector is an important issue for regional analysis because the more disaggregated the data are, the better will be the potential for understanding regional growth patterns. Comparing our study to Glaeser *et al.* (1992), there is a limitation in our manufacturing industry employment database. For example, we cannot disaggregate this sector further at the regional level in order to analyse the share of steel production, electronics manufacturing, etc. in each region. This aggregation issue must be taken into account because with more disaggregated data, the results found here could still change in terms of sign and statistical significance (McCann *et al.*, 2008). Consequently, while we would ideally conduct this study at the level of disaggregated data by sub-sector within the manufacturing sector and also by municipality, we are in fact limited to focus on the whole manufacturing sector at the smallest regional level for which we have data for all of our variables of interest, namely the state level.³

The second data source is the Ministry of Transportation of Brazil. Its website provides a table of distances in kilometres between state capitals and other main Brazilian cities. We extracted information on distance from the states' capitals to São Paulo's centre, which is the largest centre of economic activity in Brazil.⁴ We used data from the Ministry of Transportation of Brazil because this institution has current distance data and also knowledge of the quality of infrastructure throughout the country, which is important for the estimation

³ The State of Tocantins has been excluded from the analysis because of missing data from 1981 to 1992.

⁴ An alternative variable is the states' infrastructure accessibility (McCann *et al.*, 2008, p. 8). However, despite acknowledging the role of infrastructure in regional growth in Brazil (Daumal & Zignago, 2010), we did not consider infrastructure due to a lack of data.

of distance between places (the actual distance between places is not necessarily the shortest distance).

Table 2 shows the nomenclature, definitions and the basic descriptive statistics (means and standard deviations) for the variables which are used in the econometric modelling that is discussed in section 5.

4. Growth in the Regions of Brazil

In line with the literature summarised in Section 2 we ask the question whether Brazilian states grow at roughly the same rate and, if not, whether the states with the initially smallest shares of manufacturing industry employment have the fastest employment growth over the 1981 to 2006 period.

Table 3 shows the average annual growth rates of manufacturing industry employment from 1981 to 2006. At the national level, this growth rate was 2.56%. Growth rates vary across macro regions and, within macro regions, across states. At the macro regional level, the Southeast, South and Northeast grew at 1.59%, 3.52%, and 2.88% respectively. These macro regions had relatively high manufacturing employment shares in 1981: 63.80%, 16.12% and 15.29% respectively. In contrast, the Centre West and North, with national manufacturing employment shares in 1981 of only 2.66% and 2.04% grew, respectively, at 5.38% and 7.46%, on average, from 1981 to 2006. At the state level, states that grew fastest are those with lower manufacturing industry employment share in 1981. For instance, in the North macro region the smallest state of Roraima grew at 11.25%. This growth rate is above those for the two biggest states in this region, Amazonas and Pará, which grew at only 4.86% and 8.92%. In the Northeast, excluding Alagoas and Sergipe states, all other small states grew faster than the three that had the largest manufacturing employment share in 1981. The growth rates for these leading states were Bahia 2.82%, Ceará 3.08%, and Pernambuco at 1.09%. In the Centre West, employment in the leading states of Goiás and Mato Grosso do Sul were surpassed by the smaller ones, Mato Grosso and Distrito Federal. The growth rates for the former two states were 5.19% and 3.69% and for the latter two 7.61% and 6.64%. In the southeast, employment in São Paulo and Rio de Janeiro grew less than in Minas Gerais and Espírito Santo: the growth rates for the former two states were 1.28% and -0.05%; for the latter two 3.94% and 3.99%. Lastly, the South's biggest state, Rio Grande do Sul, grew 2.77%, which was less than 4.26% and 3.96% for Paraná and Santa Catarina, respectively.

Barro & Sala-i-Martin (2004, p. 464) show how absolute convergence can be defined in relation to a given steady state. This is verified by calculating the variance (σ^2) of income, product (or output) or another variable of the researcher's interest over time. If there is over time a redistribution of employment to those regions that initially have the smallest share of employment, there is absolute convergence, at least in terms of employment. Assuming that there is a common steady-state distribution of manufacturing industry employment across Brazilian states, Figure 1 suggests such absolute convergence across states with a downward trend in the variance of the logarithm of manufacturing industry employment from 3.68 in

1983 to 1.95 in 2006, towards its unknown steady-state.⁵ So, in line with the previously outlined literature that demonstrated real income convergence in Brazil, we see that there has also been manufacturing industry employment convergence. A descriptive discussion of the national and regional economic forces that led to this outcome over the 1981-2006 period can be found in Matlaba (2012, chapters 2 and 3).

Table 3: Evolution of Manufacturing Industry Employment in Brazil, 1981-2006

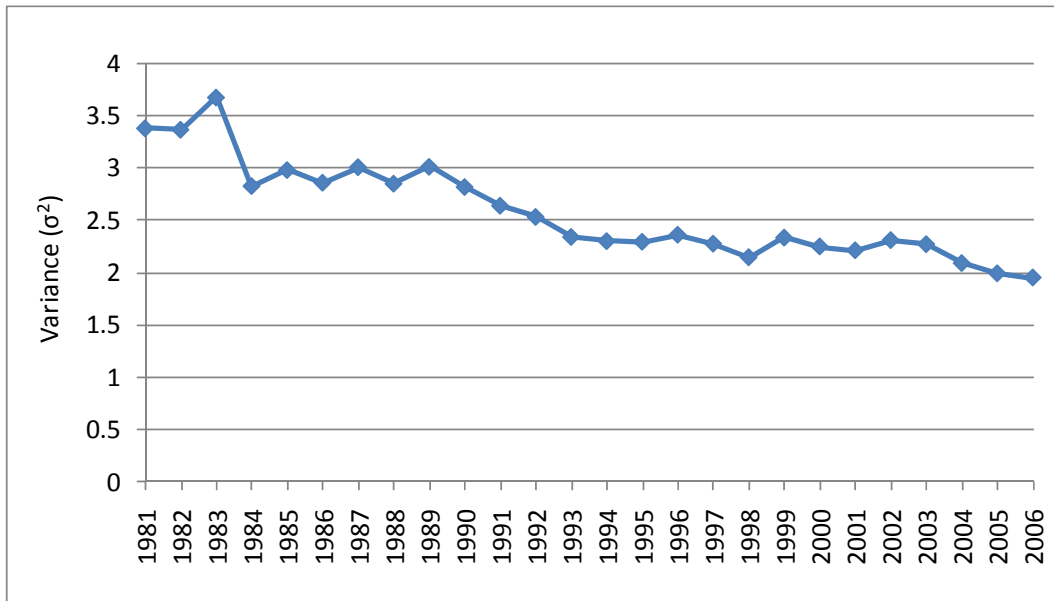
Macro-Region	State	SSR81	SSC81	SSR06	SSC06	AVAMEGS	AVAMEGR
North	Acre	3.22%	0.07%	2.97%	0.21%	7.13%	
	Amazonas	43.65%	0.89%	22.82%	1.59%	4.86%	
	Amapá	2.19%	0.04%	1.42%	0.10%	5.73%	
	Pará	43.15%	0.88%	62.19%	4.33%	8.92%	
	Rondônia	7.27%	0.15%	9.24%	0.64%	8.42%	
	Roraima	0.53%	0.01%	1.36%	0.09%	11.25%	
			100.00%	2.04%	100.00%	6.96%	
Northeast	Alagoas	5.63%	0.86%	3.77%	0.62%	1.27%	
	Bahia	22.07%	3.37%	21.73%	3.60%	2.82%	
	Ceará	25.54%	3.90%	26.84%	4.44%	3.08%	
	Maranhão	4.81%	0.74%	8.66%	1.43%	5.23%	
	Paraíba	6.15%	0.94%	8.86%	1.47%	4.34%	
	Pernambuco	23.73%	3.63%	15.18%	2.51%	1.09%	
	Piauí	2.78%	0.43%	4.55%	0.75%	4.84%	
	Rio Grande do Norte	5.26%	0.80%	6.26%	1.04%	3.58%	
	Sergipe	4.03%	0.62%	4.15%	0.69%	3.00%	
		100.00%	15.29%	00.00%	16.55%		2.88%
Centre West	Distrito Federal	6.74%	0.18%	9.25%	0.50%	6.64%	
	Goiás	54.91%	1.46%	52.41%	2.81%	5.19%	
	Mato Grosso do Sul	26.30%	0.70%	17.27%	0.93%	3.69%	
	Mato Grosso	12.05%	0.32%	21.08%	1.13%	7.61%	
			100.00%	2.66%	100.00%	5.37%	
Southeast	Espírito Santo	1.85%	1.18%	3.37%	1.69%	3.99%	
	Minas Gerais	12.09%	7.71%	21.75%	10.89%	3.94%	
	Rio de Janeiro	18.11%	11.56%	12.01%	6.01%	-0.05%	
	São Paulo	67.94%	43.35%	62.87%	31.49%	1.28%	
			100.00%	63.80%	00.00%	56.08%	
South	Paraná	26.24%	4.23%	31.55%	6.47%	4.26%	
	Rio Grande do Sul	48.01%	7.74%	39.76%	8.15%	2.77%	
	Santa Catarina	25.76%	4.15%	28.70%	5.89%	3.96%	
			100.00%	16.12%	00.00%	20.51%	
	Brazil		99.91%		99.46%	2.56%	2.56%

Notes: SSR81 (SSR06) and SSC81 (SSC06) are States' Manufacturing Industry Employment Regional Shares and States' Manufacturing Industry Employment Country Shares in 1981 (2006). AVAMEGS=Annual Average of Manufacturing Industry Employment Growth at the State level. AVAMEGR=Annual Average of Manufacturing Industry Employment Growth at the Macro Region level. The total values of the second and fourth columns are less than 100% due to data deficiencies discussed in the text.

Source: IPEA

⁵ However, within macro-regions (for which figures are not shown here) the variance in the North and Southeast is higher than that in the Northeast, Centre-West and South. Only the variance of the Southeast and South shows a clear decline. This suggests that the macro-regions behave differently. We return to this issue in our conclusion.

Figure 1: Variance of Logarithm of Manufacturing Employment Across Brazilian States



5. Empirical Results and Discussion

Cross-section Models⁶

Table 4 shows results of two cross-section models. In the Gleaser *et al.*-type model the dependent variable is 1985-95 growth of manufacturing industry employment. Specialisation and competition are defined in the same manner as in Glaeser *et al.* (1992). Diversity is defined as in Combes' (2000) study that applied Glaeser *et al.*'s model using data on manufacturing and services in the regions of France. Distance is defined as in Henderson *et al.*'s (1995) study of industry employment in cities of the United States. Education is measured by the states' literacy rate.

⁶ The test for multicollinearity in all estimated models in this paper indicates that multicollinearity is not a concern. According to McCann *et al.* (2008, p. 10), multicollinearity is not present if the Variance Inflation Factor (VIF) is smaller than 10. The calculated VIF is relatively low. For cross-section models, the average VIF ranges from 2.11 to 4.09, and the highest VIF ranges from 2.72 to 6.52 associated to the competition and diversity variables, respectively. For the remaining models, the VIF values are much lower. For panel models with annual data, the average VIF ranges from 1.21 to 2.48, and the highest VIF ranges from 1.29 to 7.14 associated with distance and logarithm of specialisation variables, respectively, when excluding an-"outlier" case of VIF equals to 12.32 associated with the distance variable. For pooled-periods cross-section models, the average VIF ranges from 1.24 to 3.03, and the highest VIF ranges from 1.38 to 7.38 associated to specialisation and logarithm of specialisation variables, respectively, again, excluding an-"outlier" case of VIF equal to 9.61 associated with distance. Another way to test for multicollinearity is to correlate explanatory variables. Under this approach, multicollinearity is a concern if the correlation of a pair of variables is above the threshold absolute value of 0.80 (McCann *et al.*, 2008). The regressors correlation across this paper's cross-section models ranges from 0.0004 (between diversity and specialisation) to 0.6897 (between diversity and competition). On the other hand, correlation across variables of both panel models with annual data and pooled-periods cross-section models ranges from 0.39 (between distance and specialisation) to 0.55 (between logarithm of literacy and logarithm of diversity). In this latter range, we excluded an-"outlier" case of absolute correlation of 0.90 (between São Paulo state's dummy and logarithm of distance) for the pooled-periods models.

Table 4: Cross-section Models

Variables	Glaeser et al. Model	Combes Model
Constant	-1.1089*** (-2.83)	-2.8412** (-2.13)
Specialisation	-0.1802 (-1.11)	-0.0257 (-0.15)
Competition	0.2496** (2.22)	0.3801*** (2.60)
Diversity	-0.1730 (-0.79)	0.1227 (0.85)
Distance	0.0001 (1.66)	0.0365*** (2.35)
Literacy	0.0150*** (2.85)	0.6166** (2.06)
Number of observations	26	26
R^2	0.7524	0.7306

Significance levels: *=10%, **=5%, and ***=1%.

Notes: Dependent Variable: 1985-95 Manufacturing Industry Employment Growth. Estimations are with robust standard errors. While the dependent variable in the left column is VDEP (see Table 2), in the right column it is VDEP2. Values of t statistics are in brackets. The left column also differs from the right column in that in the left column the natural logarithm has been taken for all independent variables, as in Combes (2000).

The models have been estimated by Ordinary Least Squares (OLS). In terms of the significant findings in the Glaeser et al. model, it can be noted that the coefficient of competition is positive and statistically significant. This result corroborates with both Porter and Jacobs' externalities. The coefficient on distance is insignificantly different from zero suggesting independence between distance from the largest market of São Paulo and growth.⁷ Finally, the coefficient of the literacy rate is positive as expected and also statistically significant. Hence spatial variation in the literacy rate is positively correlated with employment growth.

The Combes model in Table 4 refers to the specification of Combes (2000). The dependent variable is manufacturing industry employment growth, VDEP2. The independent variables are similar to those of the Glaeser et al. model; however, natural logarithms have been taken. The results indicate that the coefficient of competition is again positive and statistically significant, supporting both Porter and Jacobs' externalities. Moreover, the coefficients of distance from São Paulo and the literacy rate are also positive and statistically significant.

⁷ In many of our models in this paper the coefficient of distance is either positive but not statistically significant or positive and statistically significant. This is against the expectation of a negative relationship with distance from states' to São Paulo's capital and economic growth. However, states that are far from São Paulo may have their growth more dependent on their neighbours' states, rather than on the São Paulo state market itself. This is consistent with the notion of 'clubs of convergence' (a group of states that converge within their macro region) stated in the conclusion section.

All cross-section models have high explanatory power: the adjusted R^2 are 0.7524 and 0.7306 respectively. However, there are only 26 observations for each regression equation. Our cross-section models are comparable with previous studies. We consider the findings that are simultaneously significant in the present paper and in the previous seminal papers by Glaeser *et al.* (1992) for the USA and by Combes (2000) for France. Glaeser *et al.* (1992) estimate three regression models in city-industry in which the dependent variable is 1956-1987 employment growth, 1956-1987 wage growth, and the 1956-1987 employment growth after excluding the four biggest industries.

We compare our results with those of Glaeser *et al.*'s first regression model. For this model these authors estimated four equations with specialisation in the 1st and 4th estimations and competition in the 2nd and 4th. They found a coefficient of specialisation that is, respectively, -0.0128 and -0.00799. This supports Jacobs' externalities. This result is rejected for Brazil because our specialisation coefficient in Table 4 is negative but not statistically significant. For competition: Glaeser *et al.* (1992) found 0.587 and 0.561 which is consistent with Porter and Jacobs' hypothesis. Although our values are less than half, we find qualitatively similar results.

Although Combes (2000) applied Tobit (for externalities indicators) and Probit (for regional dummies and density of employment) estimation methods rather than OLS, to address the problem of truncated data because his sample involved French plants of more than 20 workers, a comparison between Combes' results and ours can still be made. This is because Combes also compared his results to those of Glaeser *et al.* who employed OLS instead of Probit and Tobit methods. Combes presents Global Regressions Estimations (Table 1, p. 340) and Annual Global Regressions (Table 4, p. 352) for both manufacturing employment and services. We compare his finding for manufacturing employment with our estimations.

Combes found -0.088 for specialisation on Global Regressions (p. 340) and -0.033 for specialisation on Annual Global Regressions (p. 352). We also got a negative coefficient for specialisation in Brazil in the estimates of the right column of Table 4, but it is statistically insignificant. For competition, Combes found for the aforementioned regressions -0.154 and -0.013, respectively whereas we find 0.3801, which supports both Porter and Jacobs' externalities. However, Combes used the *inverse* of our competition variable, which he named the size of plants. The expression is, "size_{z,s} = $\frac{\text{emp}_{z,s}/\text{nbr}_{z,s}}{\text{emp}_s/\text{nbr}_s}$, where $\text{nbr}_{z,s}$ and nbr_s are the number of plants belonging to sector s in ZE z and France, respectively" (Combes, 2000, p. 337). Consequently, our findings reconfirm those of Combes with respect to competition. For diversity, Combes found -0.051 and -0.026 for the Global and Annual Global Regressions, respectively. These results support both MAR's and Porter's externalities. The coefficient of diversity in the right column of Table 4 is positive, but statistically insignificant.

Annual Panel Models

Estimated coefficients of four panel models with annual data are shown in Table 5.⁸ In equations 5.1 and 5.2 the dependent variable is VDEP, following Glaeser *et al.* (1992), whereas in equations 5.3 to 5.4 it is VDEP2, following Combes (2000). The role of education in growth was addressed by inclusion of the literacy rate in equations 5.1 and 5.3 as well as its annual change in equations 5.2 and 5.4. Equations 5.2 and 5.4 differ from 5.1 and 5.3 in that that the former include year and state dummy variables. The natural logarithm is taken for the specialisation, diversity and distance variables in equations 5.3 and 5.4 and for the literacy rate in the equation 5.3.

Table 5: Panel Models with Annual Data

Variables	Eq. 5.1: OLS Model	Eq. 5.2: OLS Model with Year and State Fixed Effects	Eq. 5.3: OLS Model	Eq. 5.4: OLS Model with Year and State Fixed Effects
Constant	-0.1158** (-2.55)	-0.5912** (-2.42)	-0.0351 (-0.15)	-0.2695*** (-3.85)
Lag of the dependent variable	-0.2956*** (-4.46)	-0.3622*** (-8.02)	-0.3344*** (-5.21)	-0.3801*** (-7.66)
Specialisation	0.0298 (1.07)	0.4676*** (3.49)	0.0616* (1.92)	0.4370*** (4.73)
Diversity	0.0126 (0.32)	-0.0381 (-0.68)	0.0639 (1.30)	-0.0443 (-0.60)
Distance	0.0192** (2.36)	0.0001*** (4.26)	0.0161** (2.41)	0.0226*** (5.67)
Literacy rate	0.0012* (2.08)		-0.0060 (-0.11)	
Annual change in the literacy rate		-1.0081 (-1.34)		-0.4527 (-0.78)
Number of observations	624	624	624	624
Adjusted R^2	0.1029	0.3308	0.1396	0.3584

Significance levels: *=10%, **=5%, and ***=1%.

Notes: The dependent variable is manufacturing industry employment growth. Estimations are with robust standard errors. Values of t statistics are in brackets. The dependent variable in equation 5.1 and 5.2 is VDEP, in equations 5.3 to 5.4 it is VDEP2. Coefficients of year fixed effects for 1983, 1984-1985, 1987-1988, 1990-2001, 2003, 2005 and 2006 are negative and statistically significant. State dummies for Roraima and São Paulo were dropped. Coefficients of state fixed effects in equation 5.2 are positive and statistically significant for Acre, Amapa, Para, Rondonia, Bahia, Maranhão, Paraíba, Piauí, Sergipe, Distrito Federal, Goiás, Mato Grosso do Sul, Mato Grosso, Espírito Santo, Minas Gerais, Rio de Janeiro, São Paulo, Paraná and Rio Grande do Sul; those for Amazonas and Ceará are negative and statistically significant. All states' dummies coefficients in equation 5.4 are positive and statistically significant, except for Rio Grande do Sul, which had a negative and statistically significant coefficient and Santa Catarina, for which the dummy coefficient is not statistically significant. State dummies for Amazonas and São Paulo were dropped.

⁸ Table 3 shows the descriptive statistics for the variables of the cross-section, panel and pooled-periods cross-section models. These latter two groups of models allow for some control of (time invariant) omitted variables and may yield therefore more accurate externalities parameter estimates.

In all four models, the coefficient of the lag of the dependent variable is negative as expected and also statistically significant. Because the dependent variable is in growth rates, the negative coefficient on the lagged growth rate is consistent with an autoregressive process in levels. The autocorrelation coefficient is one minus the reported coefficient and is therefore around 0.6 to 0.7. In all four columns, the coefficient of specialisation is positive and statistically significant in all but the first column. Hence the results are broadly supportive of both MAR and Porter externalities. The coefficient of diversity is not statistically significant.

The coefficients on the distance variable show that employment growth in Brazil is faster the further the state is from São Paulo. In equation 5.1 the coefficient of the literacy rate is, as expected, positive and statistically significant. That is, the higher the literacy rate, the higher economic growth. However, the literacy rate is not statistically significant in models that include time fixed effects. The reason is that the literacy rate tends to only slowly change over time and has therefore now predictive power alongside a time trend. However, the rate of change in the literacy level is more variable but generally inversely related with the level of literacy, i.e. there is spatial literacy convergence. This implies that we would expect the coefficient on the annual change in the literacy rate to be negative: a high rate of change in literacy is equivalent to a low level of literacy. The latter suggests a low level of human capital that would be an impediment to growth. Nevertheless, while the coefficient of the annual change in literacy is negative as expected in columns 5.2 and 5.4, these coefficients are statistically insignificant. The explanatory power is relatively low for equations 5.1 and 5.3, with adjusted R^2 of 0.1029 and 0.1396 respectively. However, as expected, the adjusted R^2 is much higher for equations 5.2 and 5.4, which are both models with time and state dummies.

Pooled-Periods Models

This sub-section presents models of average annual manufacturing employment growth for pooled periods. The periods are 1981-1990, 1991-1998 and 1999-2006. This yields a total of 78 observations. We estimate again four models. In all models, the change in literacy is used rather than literacy rate itself. Table 6 shows the results. In equations 6.1 and 6.2 the dependent variable is based on Glaeser *et al.* (1992), whereas in equations 6.3 and 6.4 we follow Combes (2000). All models include the following explanatory variables: specialisation, diversity, distance, and change in the literacy rate. Additionally there are year and state dummies in equations 6.2 and 6.4.⁹ In equations 6.3 and 6.4 the natural logarithm was taken for all independent variables, except for the change in literacy over successive periods.

The coefficient of distance is positive and statistically significant in all four columns. Again, growth is stronger the further the region is from São Paulo. The coefficient of the change in literacy is negative and statistically significant in equations 6.1 and 6.3. Both the distance effect and the change in literacy effect are consistent with neoclassical growth

⁹ We also estimated fixed, random and between-effects models. The results are available upon request. The panel models not reported here do not affect our conclusions.

theory. Since São Paulo has the highest level of income and the regions furthest away from São Paulo are the least developed, those regions will have the fastest growth rates. Additionally, the rate of literacy increases the fastest where the level of human capital per capita is the lowest and this depresses growth.

Table 6: Pooled-Periods Cross-Section Models

Variables	Eq. 6.1: OLS Model	Eq. 6.2: OLS Model with Year and State Fixed Effects	Eq. 6.3: OLS Model	Eq. 6.4: OLS Model with Year and State Fixed Effects
Constant	0.0466 (1.40)	-0.1365** (-2.01)	0.0073 (0.39)	-0.0251 (-1.37)
Specialisation	-0.0079 (-0.59)	0.1256*** (3.94)	-0.0033 (-0.25)	0.0988*** (4.32)
Diversity	0.0044 (0.12)	-0.0472 (-0.66)	-0.0039 (-0.11)	-0.0265 (-0.44)
Distance	0.0115*** (3.42)	0.0001*** (7.08)	0.0085*** (3.00)	0.0095*** (4.64)
Change in literacy over successive periods	-0.3788*** (-4.21)	0.1810 (1.09)	-0.3999*** (-3.87)	0.1952 (1.22)
Number of observations	78	78	78	78
R ²	0.1990	0.7030	0.1813	0.7086

Significance levels: *=10%, **=5%, and ***=1%.

Notes: The dependent variable is the rate of average annual manufacturing industry employment growth per period. The periods are 1981-1990, 1991-1998 and 1999-2006. Estimations are with robust standard errors. Values of *t* statistics are in brackets. In equation 6.2 the dummies for Roraima and São Paulo were dropped. Only dummy coefficients for Amazonas and Bahia are negative and statistically significant. Only dummy coefficients for Amapá, Pará, Rondonia, Distrito Federal, Goiás, Mato Grosso do Sul, Mato Grosso, Espírito Santo, Minas Gerais, Rio de Janeiro and Paraná are positive and statistically significant.

In equations 6.3 and 6.4, the natural logarithm was taken for all independent variables, except for the change in literacy. In equation 6.4, the dummies for Amazonas and São Paulo were dropped. Only dummy coefficients for Acre, Amapá, Pará, Rondonia, Roraima, Alagoas, Ceará, Distrito Federal, Goiás, Mato Grosso do Sul, Mato Grosso, Paraná, Alagoas, Paraíba, Piauí, Rio Grande do Norte, Sergipe and Minas Gerais are positive and statistically significant; those for Rio Grande do Sul and Santa Catarina are negative and statistically significant.

OLS estimation with year and state dummy variables (equations 6.2 and 6.4) finds that the coefficient on specialisation is positive and statistically significant, which confirms both MAR and Porter's externalities. The coefficient of diversity is not significant in any of the models. However, similar to McCann *et al.* (2008), we see that there is an issue of stability of sign and statistical significance. By moving from the basic cross-section models to the panel models, the results with respect to the impact of the externalities appear unstable: while cross-section models confirm competition externalities, panel and pooled models "prove" the

specialisation externality.¹⁰ With respect to competition externalities we find that these are positive and confirmed in cross-section models and thereby lend support for Porter's and Jacobs' theories. However, because of a lack of data on establishments (recall from the expression for competition in Table 3 how this depends on the number of establishments), the competition externality appears only in cross-section models. Therefore this is not comparable across models. With respect to specialisation we find that this externality was rejected in cross-section and simple pooled OLS models. However, it has been confirmed in annual panel and pooled models with year and state dummy models, and in fixed effects models. These results provide support for both MAR and Porter's theories. Finally, it can be shown that presence of diversity externalities can only be confirmed in cross-section model without the education variable (results available upon request).

6. Conclusions

Many studies have documented real income convergence among the Brazilian states. This convergence was also reinforced by this study through assessment of Brazilian states manufacturing industry employment growth which shows that poor states grew faster than rich ones from 1981 to 2006. The main question of this study is why this happened. This paper attempted to answer this question by analysing the importance of the theories of dynamic externalities proposed by Marshal-Arrow-Romer (1986), Porter (1990) and Jacobs (1969) for the explanation of that convergence using Brazil's employment data.

We found that the aforementioned theories of dynamic externalities help to explain convergence among Brazilian states. First, the estimated cross-section models confirm competition along the lines of Porter and Jacobs. Even though our findings partially contrast previous studies such as those by Glaeser *et al.* (1992) and Combes (2000), they are in accordance with other international studies summarised by De Groot *et al.* (2009). However, this partial confirmation must be judged with caution because the data used in the above studies are more detailed with respect to manufacturing industry employment than our data.

Secondly, using models not previously applied in this context by Glaeser *et al.* (1992) and by Combes (2000) in Tables 5 and 6, we confirm MAR and Porter's specialisation externalities. Thirdly, analysis of sign stability showed that while for cross-section models competition externalities under Porter and Jacobs' theories are confirmed, for the pooled OLS with year and state dummy models and fixed effects models, specialisation externalities become observed only under MAR's theory. This suggests that while MAR and Porter's (low competition or more specialisation) theories are valid in some macro-regions, Jacobs' (high competition or diversity) theory is important for the other regions. This result is consistent with the 'clubs of convergence' hypothesis of the convergence literature for Brazil (see, for example, Brauch & Monasterio, 2007; Gondim *et al.*, 2007; Coelho & Figueiredo, 2007; and Magalhães *et al.*, 2005).

¹⁰ Our panel and pooled model results are unique and non-comparable with the previous seminal papers (Glaeser *et al.*, 1992 and Combes, 2000) because both only employed cross-section data.

Finally, this study has two caveats. First, to understand Brazilian macro-regions behaviour, it is important to employ disaggregated panel models that focus on individual regions because externalities operate differently in different parts of the country. Secondly, each type of externalities appears captured by specific models: on the one hand, cross-section models picture competition externalities under both Porter and Jacobs' theories and possibly diversity also under Jacobs' theory; on the other hand, annual panel and pooled models capture specialisation externalities under MAR and Porter if those models are estimated by OLS with both year and state dummies or with fixed effects. Further research is needed on why changing the model specification from a cross-section to a panel approach leads to such apparently contradicting results. These two caveats will be investigated in future research.

References

- Abreu, M. P. (2008). The Brazilian economy, 1980-1994. In L. Bethel (Ed.), *The Cambridge history of Latin America* (Vol. IX) (pp. 395-430). Cambridge, England: Cambridge University Press.
- Abreu, M. P., & Werneck, R. L. F. (2008). The Brazilian economy, 1994-2004. In L. Bethel (Ed.), *The Cambridge history of Latin America* (Vol. X) (pp. 431-454). Cambridge, England: Cambridge University Press.
- Andrade, T. A., & Serra, R. V. (2001). *Análise do desempenho produtivo dos centros urbanos brasileiros no período 1975/96*. Available from: <http://www.nemesis.org.br/>
- Azzoni, C. R. (1997). *Distribuição pessoal de renda nos estados e desigualdade de renda entre estados no Brasil: 1960, 70, 80 e 91*. Available from: <http://www.nemesis.org.br/>
- Azzoni, C. R. (1999). Regional income inequalities by income class in Brazil: Is inequality higher among rich or poor people? *Working Paper Series*. doi: 10.2139/ssrn.202688.
- Azzoni, C. R. (2001). Economic growth and regional income inequality in Brazil. *The Annals of Regional Science*, 35, 133-152.
- Azzoni, C. R. (2003). Sobre a necessidade da política regional. *Nereus Working Paper #08*. Available from: www.econ.fea.usp.br/nereus.
- Azzoni, C. R., Menezes-Filho, N., Menezes, T., & Silveira-Neto, R. (2000). *Geography and income convergence among Brazilian states*. Retrieved July 21, 2008 from Inter-American Development Bank Research Department website: http://www.iadb.org/res/pub_desc.cfm?pub_id=R-395.
- Baer, W. (2008). *The Brazilian economy: Growth and development* (6th Ed.). Boulder, CO: Lynne Rienner.
- Barossi-Filho, M., & Azzoni, C. R. (2003). A time series analysis of regional income convergence in Brazil. *Nereus Working Paper #09*. Available from: www.econ.fea.usp.br/nereus.
- Barro, R. J., & Sala-i-Martin, X. (2004). *Economic growth*. Cambridge, MA: MIT Press.
- Brauch, M. D., & Monasterio, L. M. (2007). *Spatial dynamics of economic growth in Brazil (1970-2000): An empirical approach*. Available from: <http://www.nemesis.org.br/>
- Coelho, R. L. P., & Figueiredo, L. (2007). Uma análise da hipótese da convergência para os municípios Brasileiros. *Revista Brasileira de Economia*, 61(3), 331-352.
- Combes, P. P. (2000). Economic structure and local growth: France, 1984–1993. *Journal of Urban Economics*, 47, 329-355.
- CORECON - Conselho Regional de Economia de São Paulo (or, São Paulo's Regional Council of Economics): http://www.coreconsp.org.br/indicadores/dip/planos_economicos.html.
- Daumal, M., & Zignago, S. (2010). Measure and determinants of border effects of Brazilian states. *Papers in Regional Science*, 89(4). doi:10.1111/j.1435-5957.2009.00265.x
- De Groot, H. L. F., Poot, J., & Smit, M. J. (2009). Agglomeration externalities, innovation and regional growth: theoretical perspectives and meta-analysis. In R. Capello & P. Nijkamp (Eds.), *Handbook of regional growth and development theories* (pp. 256-281). Cheltenham, England: Edward Elgar Publishing.

- Ferreira, A. (2000). Convergence in Brazil: Recent trends and long-run prospects. *Applied Economics*, 32, 479-489.
- Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities. *Journal of Political Economy*, 100(6), 1126-1152.
- Gomes, G. M. (2002). *Regional development strategies in Brazil*. International Conference on Regional Development and Foreign Direct Investment in Fortaleza, Brazil. Available from: <http://www.oecd.org>.
- Gondim, J. L. B., Barreto, F. A., & Carvalho, J. R. (2007). Condicionantes de clubes de convergência no Brasil. *Estudos Econômicos*, 37(1), 71-100.
- Henderson, V., Kuncoro, A., & Turner, M. (1995). Industrial development in cities. *Journal of Political Economy*, 103(5), 1067-1090.
- IPEA - Instituto de Pesquisa Econômica Aplicada: www.ipea.gov.br. (Accessed in the first semester of 2009).
- Jacobs, J. (1969). *The Economy of Cities*. New York: Vintage.
- Macedo, R., & Barbosa, F. (1997). Brazil: Instability and macroeconomic policies. In M. J. F. Willumsen & E. G. da Fonseca (Eds.), *The Brazilian economy: Structure and performance in recent decades* (pp. 1-23). Miami, FL: North-South Center Press at the University of Miami.
- Magalhães, A., Hewings, G. J. D., & Azzoni, C. R. (2005). Spatial dependence and regional convergence in Brazil. *Investigaciones Regionales*, 6, 5-20.
- Matlaba, V. J. (2012) *Regional Socio-Economic Transformation in Brazil*. PhD Dissertation. Hamilton: University of Waikato.
- McCann, P., Mameli, F. M., & Faggian, A. (2008). Employment growth in Italian local labour systems: Issues of model specification and sectoral aggregation. *Spatial Economics Analysis*, 3(3), 1-18.
- Mossi, M. B., Aroca, P., Fernández, I. J., & Azzoni, C. R. (2003). Growth dynamics and space in Brazil. *International Regional Science Review*, 26(3), 393-418.
- Porter, M. (1990). *The competitive advantage of nations*. New York: Free Press.
- Rangel, L. A., Andrade, J., & Divino, J. A. (2008). *Economic growth and income inequality in Brazil: Analyzing the comparable minimum areas*. Retrieved March 21, 2009, from European Economics and Finance Society Prague Papers website: www.eefs.eu/conf/Prague/Papers/S602_Andrade_Divino_Rangel.pdf.
- Resende, G. M. (2011). Multiple dimensions of regional economic growth: The Brazilian case, 1991-2000. *Papers in Regional Science*, 90(3). doi:10.1111/j.1435-5957.2010.00336.x
- Road Transportation Information website of The Ministry of Transportation of Brazil: <http://www.transportes.gov.br/bit/inrodo.htm>. <Accessed March, 2009>.
- Rolim, C. (2008). *Decomposing the GDP growth of the great Brazilian regions: Dynamic shift-share on recent Brazilian economic cycles*. World Congress of Regional Science Association International (March 16th-19th, 2008), University of São Paulo, Brazil. Available from: <http://www.aber.fea.usp.br/rsai2008/new/html/event/event.asp>.
- Romer, P. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, 94(5), 1002-1037.
- Silva, M. V. B., & Silveira-Neto, R. M. (2007). Determinantes da localização industrial no Brasil e geografia econômica: Evidências para o período pós-real. *Revista Economia*, 8, 269-288.
- Silveira-Neto, R., & Azzoni, C. R. (2006). Location and regional income disparity dynamics: The Brazilian case. *Papers in Regional Science*, 85(4), 599-613.
- Wooldridge, J. M. (2002). *Solutions manual and supplementary materials for econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.
- Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

Table 2: Variables used in the analysis

Dependent Variables			
Variable	Expression	Explanation	Means and standard deviations
Growth of Employment	$VDEP = \ln\left(\frac{E_t}{E_{t-1}}\right)$ (as in Glaeser <i>et al.</i> (1992))	States' manufacturing employment growth from $t-1$ to t .	$t=1995$; $t-1=1985$: 0.2898 (0.3260) annual t from 1981 to 2006: 0.0457 (0.1726)
Growth of Employment	$VDEP2 = y_{z,s} = \ln\left(\frac{emp_{z,s,t}}{emp_{z,s,t-1}}\right) - \ln\left(\frac{emp_{s,t}}{emp_{s,t-1}}\right)$ (as in Combes (2000))	The difference between the state z and national manufacturing employment s growth from $t-1$ to t .	$t=1995$; $t-1=1985$: 0.2187 (0.3260) annual t from 1981 to 2006: 0.0201 (0.1660)
Annual Average Growth of Employment	$Avgrwth_1 = \ln(E_{1990}/E_{1981})/9$ $Avgrwth_2 = \ln(E_{1998}/E_{1990})/8$ $Avgrwth_3 = \ln(E_{2006}/E_{1998})/8$	The average annual growth of manufacturing employment, E , in each period for each state.	Pooled: 0.0453 (0.0476)
Independent Variables			
Specialisation	$Spec = \frac{E_{z,s}/E_z}{E_s/E}$ (as in Glaeser <i>et al.</i> (1992))	The ratio between state z manufacturing employment s share and Brazil's manufacturing employment s share.	1985: 0.7153(0.3626) Annual 1981 to 2006: 0.7535 (0.3474) Pooled: 0.7664 (0.3372)
Diversity	$Div_{z,s} = \frac{1/\sum_{s' \neq s}^S \left[\frac{emp_{z,s'}}{emp_z - emp_{z,s}}\right]^2}{1/\sum_{s' \neq s}^S \left[\frac{emp_{s'}}{emp - emp_s}\right]^2}$ (as in Combes (2000))	Diversity is the inverse of an Herfindahl index of sectoral concentration based on the share of all sectors, except the one considered, where S is the total number of sectors and z is state. The nine sectors s are: agriculture & fishing; commerce; construction; electricity, water & gas; finance; manufacturing; mining; services; and transportation & communications.	1985: 1.0048 (0.4577) Annual 1981 to 2006: 0.9411 (0.2181) Pooled: 0.9205 (0.1272)

Competition	$\text{Comp} = \frac{N_{z,s}/E_{z,s}}{N_s/E_s}$ (as in Glaeser <i>et al.</i> (1992))	The ratio of state z manufacturing establishment s to state z manufacturing employment s . This ratio is divided by its national equivalent in year t .	1985: 1.5799 (0.7378)
Lag of the Dependent Variable	$\text{VDEP}_{t-1} = \left[\ln \left(\frac{E_t}{E_{t-1}} \right) \right]_{t-1}$	The growth of manufacturing industry employment in the previous period in each year t , thereby starting in 1983.	Annual 1981 to 2006: 0.0470 (0.1748)
Lag of the Dependent Variable	$\text{VDEP2}_{t-1} = (y_{z,s})_{t-1}$	The difference between the state z and national manufacturing employment s growth from $t-1$ to t for the previous period in each year t , thereby starting in 1983.	Annual 1981 to 2006: 0.0208 (0.1681)
Literacy	Education = LIT _{t}	The percentage of literate people of 15 years old or more.	1985: 76 (14.08) Annual 1981 to 2006: 80.73 (11.81)
Change in Literacy	$\text{CHANGELIT}_{t-1,t} = \ln(\text{LIT}_t / \text{LIT}_{t-1})$	The change of LIT between $t-1$ and t .	1985: 0.0864 (0.0768) Annual 1981 to 2006: 0.0073 (0.0211) Pooled: 0.0607 (0.0458)
Distance	Distance = D _{i,SP} (as in Henderson <i>et al.</i> (1992))	The distance in kilometres, which is constant over the years in each state, from centres of each state i 's capital to São Paulo's city centre (the major country's market) ¹¹ .	2010: 2065 (1234.49)

Note: *Values of standard deviations are in brackets.

¹¹ The distance is based on the shortest route with the 2010 Federal, State and Municipal roads infrastructure (for São Paulo state itself distance is set equal to 1 km). We are assuming that the infrastructure was similar at the beginning of the period considered compared with the end. This assumption is important especially for annual models because these models deal with a long period in which Brazil moved through a process of long-run development.